Review of the Draft NTP SAN Trimer Technical Report (TR-573)

Prepared by Dr. Joseph K. Haseman at the Request of the SAN Trimer Association
Submitted to the National Toxicology Program, January 12, 2011

During my 33 years at the NIEHS, I was the biostatistician included in the core group of NTP scientists given the responsibility of interpreting experimental results and preparing NTP Technical Reports that summarized NTP's rodent cancer bioassays. Because of this experience, the SAN Trimer Association asked me to review the Draft NTP SAN Trimer Technical Report with special emphasis on the CNS (brain and spinal cord) tumors observed in male rats. In particular, I was asked to give an opinion regarding the appropriate level of evidence of carcinogenic activity in this study.

After examining the data from the SAN Trimer study (and the level of evidence calls made in previous NTP studies for similar patterns of brain tumor occurrence), it is my opinion that the proper call for male rats is "no evidence of carcinogenic activity." My opinion is based on the collective weight of evidence from a variety of factors as discussed below.

I. Lack of statistical significance; all brain tumor incidences are low and fall within the historical control range

None of the brain tumor increases in the SAN Trimer study are statistically significant, even by a trend test. With regard to historical control data, the NTP SAN Trimer TR cites Sills et al. (1999) as an important reference for the NTP experience with brain tumors. This paper reports that brain tumor incidences in male rat control groups have been as high as 4% for both astrocytoma and for granular cell tumors. The highest incidence of brain astrocytoma seen in male rats in the SAN Trimer study is 2%, but if the one brain astrocytoma in the high dose group is combined with the one spinal cord astrocytoma, the incidence increases to 4%, still within the historical control range. The maximum incidence of granular cell tumors observed in male rats in the SAN Trimer study was only 2%, which also falls within the historical control range of 0-4%.

II. The improved survival in the high dose group increased the likelihood of spontaneous tumor development

In the NTP SAN Trimer study, 44/50 of the high dose male rats survived until the end of the study compared with only 36/50 controls. This difference is marginally significant ($p=0.071$). The unusually good survival in the high dose group increased the likelihood of spontaneous tumor development.

III. It is unlikely that the marginally increased incidence and/or severity of nerve fiber degeneration in high dose male rats is either biologically or statistically significant,

Marginally increased incidence and the poly-3 test: One factor that apparently contributed to the NTP call of "equivocal evidence" in male rats was the slight increase in the incidence of nerve fiber degeneration that was seen following an expanded histopathology review of brain and spinal cord. Table 19 states (without giving any actual statistical analysis) that the slight increase in the incidence of nerve fiber degeneration in the spinal nerve roots observed in the high dose male rat group relative to controls (43/50 vs. 34/47) is statistically significant ($p<0.05$) by a poly-3 test. However, by a simple Fisher's exact test, these two incidences are not significantly different ($p=0.08$), and a poly-3 test, which adjusts for the improved survival in the top dose group, should logically make this difference even less significant, not more significant.

The only way that the poly-3 test could produce a $p<0.05$ high dose effect for these data would be if animals that live longer actually have a reduced likelihood of developing nerve fiber degeneration compared with animals that die early. However, this would make no sense biologically, and if this pattern of response were seen, it would indicate that one of the basic underlying assumptions of the poly-3 test was violated, namely the assumption that the risk of developing a lesion increases as an animal ages.

In fact, the standard poly-3 test may not be appropriate for these data for several reasons. First, the description of the extended evaluation (e.g., "a section of the dorsal and ventral spinal nerve roots were also evaluated in the lumbar and, less commonly, the cervical and thoracic sections of the spinal cord") suggests that most animals had multiple spinal nerve roots and/or sciatic nerves examined, and some animals may have had more nerves/nerve roots examined than others. A request has been made to obtain (not as yet received) the individual animal data from the extended evaluation to determine how many nerves/nerve roots were examined for each animal. Obviously, the more nerve roots examined for a given animal, the more likely it would be to find one with some level of degeneration. The standard poly-3 test implicitly assumes that all animals have the same number of nerve/nerve roots evaluated.

.

A second reason that the standard poly-3 test may not be appropriate is that it ignores "litter effects", which the NTP considers to be important in this particular study. For this reason, the NTP statistical analysis of neoplasms took litter effects into account by using a weighted mixed effects logistic regression model (see page 46 and Table A2 of the SAN Trimer TR). However, the analysis of nerve fiber degeneration was apparently a standard poly-3 test that did not take litter effects into account. Litter effects should be taken into account in the evaluation of nerve fiber degeneration, just as it was in the statistical analysis of neoplasms.

Finally, the standard poly-3 test assumes that the risk of developing nerve fiber degeneration increases as an animal ages, a very logical assumption. This assumption would be violated, for example, if the incidence of nerve fiber degeneration in control animals is substantially higher in the non-survivors than in the survivors. Further, as

discussed above, it appears that the only possible way that the marginally increased incidence of nerve fiber degeneration in the high dose male rat group could be statistically significant ($p<0.05$) by a standard poly-3 test (as indicated in Table 19) would be if the poly-3 test itself was inappropriate, because one of the basic underlying assumptions (noted above) was violated.

The slight increase in the incidence of nerve fiber degeneration in the sciatic nerve in male rats was not reported to be statistically significant, although the Abstract Summary Table lists this increase as a non-neoplastic effect. I do not agree that this marginal, non-significant increase is related to SAN Trimer.

Apparent lack of blinded histopathology: Another factor that may have contributed to the marginally elevated incidence of nerve fiber degeneration is the apparent lack of blinding of the extended evaluation histopathology review. In my opinion, for differences in incidence (and severity) this subtle, it is important that the extended review be carried out in a blinded fashion, without knowledge of the identity of the dosed/control group(s) or any previous (unblinded) diagnoses, in order to avoid a subtle bias toward an apparent treatment effect. The NTP SAN Trimer Technical Report should clarify whether or not the extended evaluation was a blinded histopathology review.

Marginally increased severity: The NTP SAN Trimer TR Abstract states that the increased severity of nerve fiber degeneration is even more important than the increase in incidence. However, I doubt that the slight increases in severity (e.g., 1.1 – 1.2 – 1.3 – 1.3 for nerve fiber degeneration of the sciatic nerve for the control, low, mid, and high dose male rat groups respectively) is biologically important. This slight severity increase is not reported to be statistically significant either, and could easily be due to the improved survival in the dosed groups and the apparent non-blinded review, as noted earlier. It is also unclear how the severity "score" was calculated for an animal with multiple nerve roots examined (the most extreme level of degeneration? Some sort of average response?). Interestingly, the NTP Abstract Summary Table does not report the increased severity to be related to SAN Trimer, even though earlier in the Abstract, they assert that the increased severity was even more important than the increased incidence.

Finally, if the increased incidence and/or severity of nerve fiber degeneration are somehow associated with the development of brain tumors, then why did dosed female rats (which according to Table 19 showed a much greater increased incidence of sciatic nerve fiber degeneration relative to controls than the males) show no brain tumor effects?

Thus, in my opinion, the NTP has not demonstrated that there is a statistically or biologically significant SAN Trimer effect on either the incidence or severity of nerve fiber degeneration in male rats that would potentially support a call of equivocal evidence for the brain/spinal cord tumors observed.

IV. Previous NTP calls for similar patterns of brain tumor incidences have been "No evidence"

Although the composition of NTP Peer Review Panels changes over time, it is desirable to maintain consistency in the interpretation of experimental results. Thus, it is of interest to see how similar patterns of brain tumor occurrence have been interpreted in the past.

In the Discussion section of the SAN Trimer Technical Report, the NTP attempts to justify the "equivocal evidence" call in male rats by asserting that the patterns of brain tumor incidences in the SAN Trimer study were similar to those seen in the NTP studies cited by Sills et al. (1999) as showing equivocal evidence of carcinogenic activity based on marginally increased incidences of brain tumors. However, this conclusion is very misleading for several reasons: (i) if one looks closely, the patterns of tumor occurrence in the SAN Trimer study are not similar to the tumor incidences from the studies reported by Sills et al. (1999); and more importantly, (ii) the Sills et al. (1999) paper does not consider any of the NTP studies that had low brain tumor incidences and yet were regarded as showing "no evidence" of carcinogenic activity.

To address this issue, I reviewed more than 250 previous NTP rodent carcinogenicity studies, looking for patterns of brain tumor occurrence in male F344 rats that were similar to that seen in the SAN Trimer study. Importantly, in an effort to maximize the likelihood of an "apples to apples" comparison, I eliminated from consideration those studies showing other carcinogenic effects in male rats that could possibly have been a confounding factor in the interpretation of brain tumors. Thus, I focused on those studies in which the level of evidence for carcinogenic activity in male rats was either "no evidence" or "equivocal evidence," the two possible calls for the SAN Trimer study. The results are summarized in the table below. Note that the NTP combines astrocytomas, gliomas, and/or oligodendrogliomas when evaluating glial neoplasms.

Chemical	TR	Brain tumor	Tumor incidence				Call
			Control	Low	Mid	High	
1,2,3-Benzotriazole	088	Glioma/Oligodendroglioma	0/46	3/44	-	0/46	Equivocal
Diphenhydramine HCL	355	Astrocytoma/gliomas	1/49	0/49	-	5/50	Equivocal
Furosemide	356	Meningioma	0/50	3/50	-	0/50	Equivocal
Divinylbenzene	534	Astrocytoma/Oligodendroglioma	0/49	1/50	3/50	0/50	Equivocal
Eugenol	223	Astrocytoma/Glioma	0/40	1/50	-	2/49	No evidence
Primidone	476	Astrocytoma	0/50	0/50	0/49	2/49	No evidence
CI Pigment Red 23	411	Astrocytoma/Oligodendroglioma	0/50	1/3	1/4	2/50	No evidence
Tetracycline HCL	344	Astrocytoma/Glioma*	0/50	0/50	-	2/50	No evidence
4-Methylimidazole	535	Astrocytoma/Glioma	0/50	0/50	1/50	2/50	No evidence
SAN Trimer	573	Astrocytoma *	0/50	0/50	1/50	2/50	???

* includes spinal cord

As the table shows, the four "equivocal evidence" studies produced patterns of tumor occurrence quite different from the incidences seen in the SAN Trimer study, while the five "no evidence" studies (Eugenol, Primidone, CI Pigment Red 23, Tetracycline HCL and 4-Methylimidazole) provided a very close match. In fact, the tumor incidences in the SAN Trimer and 4-Methylimidazole studies are identical. Note also that the apparent

"threshold" of tumor occurrence leading to an equivocal call seems to be three or more brain tumors in a dosed group (see the table above).

Moreover, in two of the five "no evidence" studies, there are some very interesting additional results that are not evident by simply examining the tumor incidences shown above. First, in the primidone study, survival was poor, and no high dose males survived until the end of the study. This compares with 44/50 survivors in the high dose SAN Trimer group, as noted earlier. This makes the two astrocytomas observed in the high dose primidone group even more "biologically significant" than the two observed in the high dose SAN Trimer group, in the sense that the animals in the primidone study had much less time to develop tumors. Even so, the primidone study was deemed negative.

Secondly, in the CI Pigment Red 23 study, it is noteworthy that the NTP did not even consider the high dose incidence of 2/50 brain tumors worthy of carrying out complete histopathology for the low and mid dose groups, which in theory could have found even more brain tumors, especially since brain tumors were seen in some of the few animals that were examined in these groups. The slight increase in brain tumors in the high dose group was simply dismissed out of hand.

The SAN Trimer study also showed a very low incidence of granular cell tumors of the brain/spinal cord (0/50 – 1/50 – 1/50 – 1/50) in male rats. However, these tumors are of a completely different cell type than the astrocytomas/gliomas/oligodendrogliomas, and thus they should be considered separately. This single-occurrence pattern of granular cell tumors in the dosed SAN Trimer male rat groups is even less impressive than the brain tumor incidences in the five chemicals discussed above that the NTP considered to provide no evidence of carcinogenic activity. Thus, in my opinion, the granular cell tumors likewise clearly represent no evidence of carcinogenic activity.

Moreover, in my opinion, two "no evidence" calls for biologically unrelated tumors in the brain/spinal cord do not collectively indicate an "equivocal evidence" call.

V. The extended histopathology evaluation did not find additional tumors of consequence

This is arguably the most important finding of all. In the SAN Trimer study, the NTP went to great lengths to examine additional samples of brain and spinal cord tissue in an effort to find more brain/spinal cord tumors. While all 200 male rats had brain examined in the initial histopathology review, only two male rats had spinal cords examined microscopically. Thus, the NTP carried out a more extensive evaluation that included evaluating spinal cords for all 200 male rats (apparently multiple sections per animal, which would result in hundreds or even thousands of new sections examined) and, for each animal, nine additional sections of brain (i.e., 1800 additional brain sections).

The result of this extensive re-evaluation in male rats was to find one additional spinal cord tumor: a granular cell tumor in a low dose group animal. No additional brain

tumors were found. If this brain/spinal cord tumor effect had been real (or even "equivocal"), I would have expected this additional histopathology evaluation to have found additional brain or spinal cord tumors in the mid or high dosed group. Clearly, the extended histopathology evaluation supports a "no evidence" call for male rats.

Thus, for all of these reasons, it is my opinion that "no evidence" of carcinogenic activity is the proper call for male rats in the NTP SAN Trimer study.

Joseph K. Haseman
J.K. Haseman Consulting

REFERENCE

Sills, R.C., Hailey, J.R., Neal, J., Boorman, G.A., Haseman, J.K. and Melnick, R.L. Examination of low incidence brain tumor response in F344 rats following chemical exposures in National Toxicology Program carcinogenicity studies. Toxicologic Pathology 27: 589-599, 1999.